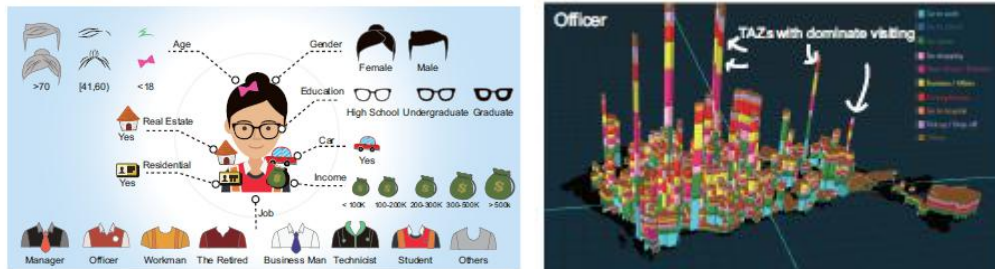


This week:

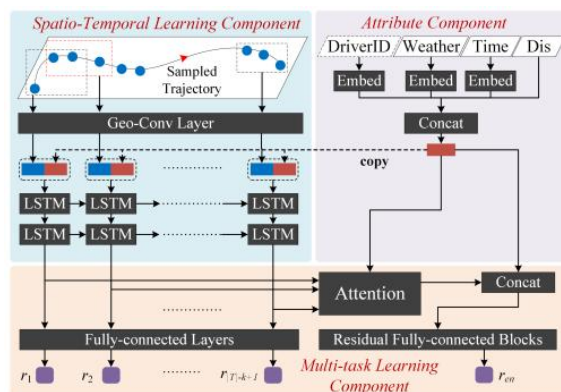
Review and read paper:

< Unfolding the City: Exploring Mobility Patterns with Individual Characteristics >



本文介绍了一种通过用户标记的 OD 数据来分析城市人群分布的可视化设计。通过卡通的视觉编码，用户能够选择不同的人群，然后分析他们的轨迹，包括他们家的分布，工作地点的分布和工作娱乐的分布。但是文章实现的方法大都是基本的统计数据 and 图表，过滤及数据处理也都是简单的操作。

< When Will You Arrive? Estimating Travel Time Based on Deep Neural Networks >



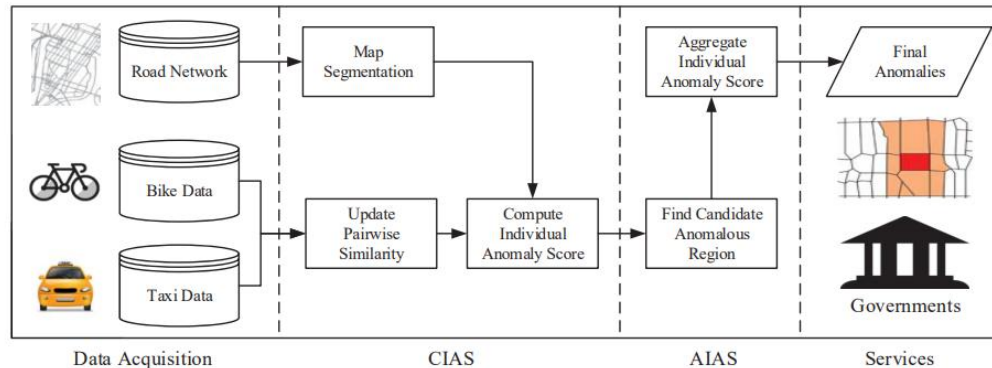
本文通过结合轨迹数据和其他信息例如天气，时间，用户习惯等其他因素，实现了一种错误率比其他方法明显低的行驶时间估计的方法。本文考虑了车辆在道路上行驶的方式（直行，左转，上下高速，右转，掉头等）。本文的模型使用了两部分，一部分通过模拟每一个路段的行驶时间，然后加起来，另一部分直接模拟整个行驶时间，这两部分都有利有弊，前者数据多但忽略了路口的情况，后者对于长距离轨迹无法预测，因为数据少。本模型能够通过模拟然后学习出两者结合的加权参数。同时如上图，对于分段预测，把整个轨迹按照数据点跨度分为很多个 slides，然后通过降维加上行驶距离然后算出每一段的估计时间模块，合起来估计整个行驶时间。同时有一个模块负责分析其他因素对于轨迹行驶时间的影响，最后的方法效果极好。本文中有把类别数据转换成向量，并衡量他们之间距离的 embedding 方法。

< Discovering Functional Zones in a City Using Human Movements and Points of Interest >



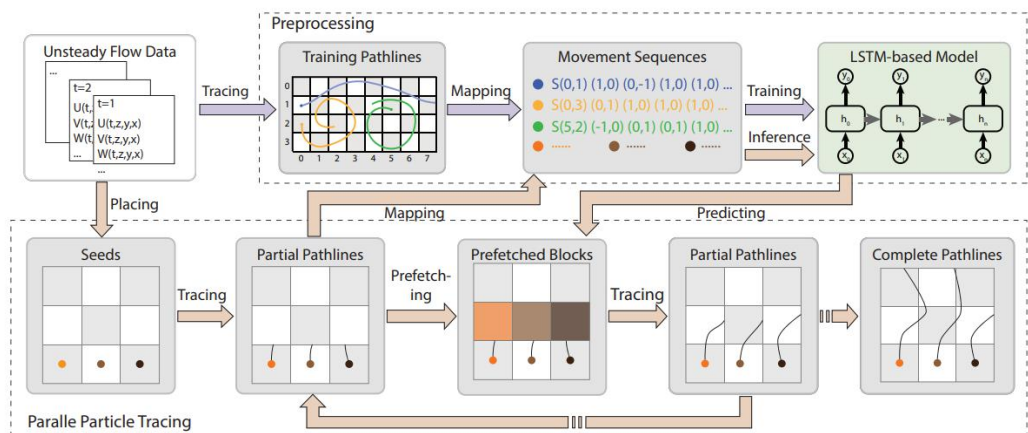
本文介绍了一种通过轨迹和 POI 数据对城市区域划分的方法。首先通过路网数据通过图像处理的方式，扩大道路线条，得到合理的区域划分，然后简化线得到精确区域划分的线。然后将轨迹数据针对每一个区域处理成为，离开该区域的时空立方体和到达该区域的时空立方体，加上该区域的 POI 类型，形成该区域的文档，然后通过 Topicmodel 进行主题抽取，抽取后对区域的 topic 分布向量集进行聚类，将区域划分成为不同的类型。本文的方法和我们的区别是区域划分使用路网，然后结合了轨迹数据进行划分。基本思路是来到同一个地方的很多区域，类型可能是一样的。同理，从同一个地方去的区域类型也很有可能是一样的。此外，该文章中有将 POI 类型进行分析时的 TF-IDF 算法。聚类的方式也是有自动选择聚类种类数量的 k-means 算法。区域聚类后的标记考虑了 4 点：1，POI 的密度，用 POI 数量除以区域面积。然后用 svd 得到 POI 参数。2，最多的移动 pattern。3，区域的核密度，用轨迹计算得出。4，用户直接标记的区域。本文处理好数据大约用了 1400 分钟，一天左右。2011 一年的出租车和地铁数据。

< Detecting Urban Anomalies Using Multiple Spatio-Temporal Data Sources>



本文应用多个数据源进行城市异常的检测，首先多个数据源将被单独计算出一个比对历史数据在区域和时间上的分数，然后通过 one-class Support Vector Machine 来发现数据中的稀有模式，然后进行最终的标记。能够高效的利用多源数据进行城市异常的检测。

< Access Pattern Learning with Long Short-Term Memory for Parallel Particle Tracing>



这项工作提出了一种基于深度神经网络的流场可视化并行粒子追踪的新型访问模式估计方法。开发了一个基于 Long Short-term Memory (LSTM) 的模型，该模型能够通过少量训练样本学习准确的访问模式，并以较小的存储开销表示学习模式。并行粒子追踪框架使用 CPU 和 GPU 进行粒子追踪任务并设计了三种不同流量数据集的各种流动可视化。

审稿	完成	
读书	阅读概率方法的前三章	
读文章		
Voice project		Ross 和师姐去开会了，下周讨论